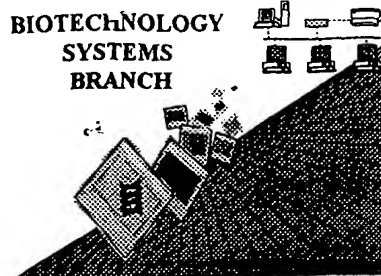


0590
0622



RAW SEQUENCE LISTING ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 09/876,796

Source: 01PE

Date Processed by STIC: 6/27/2001

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,**
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY**

FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.

PATENTIN 2.1 e-mail help: patin21help@uspto.gov or phone 703-306-4119 (R. Wax)

PATENTIN 3.0 e-mail help: patin3help@uspto.gov or phone 703-306-4119 (R. Wax)

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER VERSION 3.0 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW:

Checker Version 3.0

The Checker Version 3.0 application is a state-of-the-art Windows based software program employing a logical and intuitive user-interface to check whether a sequence listing is in compliance with format and content rules. Checker Version 3.0 works for sequence listings generated for the original version of 37 CFR §§1.821 – 1.825 effective October 1, 1990 (old rules) and the revised version (new rules) effective July 1, 1998 as well as World Intellectual Property Organization (WIPO) Standard ST.25.

Checker Version 3.0 replaces the previous DOS-based version of Checker, and is Y2K-compliant. Checker allows public users to check sequence listings in Computer Readable form (CRF) before submitting them to the United States Patent and Trademark Office (USPTO). Use of Checker prior to filing the sequence listing is expected to result in fewer errored sequence listings, thus saving time and money.

Checker Version 3.0 can be down loaded from the USPTO website at the following address:

<http://www.uspto.gov/web/offices/pac/checker>

OIPE

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/876,796

DATE: 06/27/2001
TIME: 15:33:37

Input Set : A:\SEQLIST.txt
Output Set: N:\CRF3\06272001\I876796.raw

error throughout
Does Not Comply
Corrected Diskette Needed

SEQUENCE LISTING

C--> 3 (1) GENERAL INFORMATION:

- 4 (i) APPLICANT: Horwath, K. L., et al.
5 (ii) TITLE OF INVENTION: Nucleic Acids Sequences Encoding Type III
6 Tenebrio Antifreeze Proteins and Method for Assaying

Activity.

- 7 (iii) NUMBER OF SEQUENCES: 48

C--> 8 (iv) CORRESPONDENCE ADDRESS:

- 9 (A) ADDRESSEE: Dr. Kathleen L. Horwath
10 (B) STREET: Department of Biological Sciences, Binghamton University
11 (C) CITY: Binghamton
12 (D) STATE: New York
C--> 13 (F) ZIP: 13902-6000

C--> 14 (v) COMPUTER READABLE FORM:

- 15 (A) MEDIUM TYPE: Diskette, 3.50 inch, 1.4 Mb storage
16 (B) COMPUTER: IBM AT/ATX compatible
17 (C) OPERATING SYSTEM: Windows 95/98
18 (D) SOFTWARE: Microsoft Word

19 (vi) CURRENT APPLICATION DATA:

- C--> 20 (A) APPLICATION NUMBER: US/09/876,796
C--> 21 (B) FILING DATE: 07-Jun-2001
22 (C) CLASSIFICATION:

C--> 23 (vii) PRIOR APPLICATION DATA:

- 24 (A) APPLICATION NUMBER: 60210446
25 (B) FILING DATE: June 8, 2000

C--> 26 (viii) ATTORNEY/AGENT INFORMATION:

- 27 (A) NAME: Mark Levy, Attorney-at-Law
28 (B) REGISTRATION NUMBER: 29,188
29 (C) REFERENCE/DOCKET NUMBER: RB125

C--> 30 (ix) TELECOMMUNICATION INFORMATION:

- 31 (A) TELEPHONE: 607-722-660
32 (B) TELEFAX: 607-724-2207

ERRORED SEQUENCES

C--> 60 (2) INFORMATION FOR SEQ ID NO: 2

61 (i) SEQUENCE CHARACTERISTICS:

- 62 (A) LENGTH: 566 base pairs
63 (B) TYPE: nucleic acid
64 (C) STRANDEDNESS: double
65 (D) TOPOLOGY: linear

66 (ii) MOLECULE TYPE: cDNA to mRNA

67 (iii) HYPOTHETICAL: no

68 (iv) ANTI-SENSE: no

69 (vi) ORIGINAL SOURCE:

- 70 (A) ORGANISM: Tenebrio molitor

C--> 71 (C) INDIVIDUAL ISOLATE: none

FYI! all
U.S. applications
filed on or after
July 1, 1998, and
which cannot
claim a prior
application filed
before July 1, 1998,
must be in
new sequence
Rule format.

RAW SEQUENCE LISTING

DATE: 06/27/2001

PATENT APPLICATION: US/09/876,796

TIME: 15:33:37

Input Set : A:\SEQLIST.txt

Output Set: N:\CRF3\06272001\I876796.raw

C--> 72 (G) CELL TYPE: fat body and whole organism
 73 (vii) IMMEDIATE SOURCE:
 74 (A) LIBRARY: cDNA
 75 (B) CLONE: 13.17
 76 ~~(ix) FEATURES~~
 78 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
 E--> 80 GTGGATCCAA AGAATTCCGC ACGAGACTAC TAAG ATG AAG TTG CTC
 81 Met Lys Leu Leu
 82 -15
 E--> 84 TGT TGT CTA ATC TCC CTC ATT CTG TTG GTC ACA GTT CAG GCC CTG
 85 Cys Cys Leu Ile Ser Leu Ile Leu Leu Val Thr Val Gln Ala Leu
 86 -10 -5 1
 E--> 88 ACC GAG GCA CAA ATT GAG AAA CTG AAC AAG ATC AGC AAA AAA TGT
 89 Thr Glu Ala Gln Ile Glu Lys Leu Asn Lys Ile Ser Lys Lys Cys
 90 5 10 15
 E--> 92 CAA AAT GAA AGT GGA GTG TCG CAA GAG ATC ATA ACC AAA GCT CGC
 93 Gln Asn Glu Ser Gly Val Ser Gln Glu Ile Ile Thr Lys Ala Arg
 94 20 25 30
 E--> 96 AAC GGT GAC TGG GAG GAC GAT CCT AAA CTG AAA CGC CAA GTT TTT
 97 Asn Gly Asp Trp Glu Asp Asp Pro Lys Leu Lys Arg Gln Val Phe
 98 35 40 45
 E--> 100 TGC GTG GCC AGG AAC GCC GGT CTG GCC ACG GAA TCG GGA GAG GTG
 101 Cys Val Ala Arg Asn Ala Gly Leu Ala Thr Glu Ser Gly Glu Val
 102 50 55 60
 E--> 104 GTG GTC GAC GTG TTG AGG GAG AAG GTG AGG AAG GTC ACT GAC AAC
 105 Val Val Asp Val Leu Arg Glu Lys Val Arg Lys Val Thr Asp Asn
 106 65 70 75
 E--> 108 GAC GAA GAA ACT GAG AAA ATC ATC AAT AAG TGC GCC GTC AAG AGA
 109 Asp Glu Glu Thr Glu Lys Ile Ile Asn Lys Cys Ala Val Lys Arg
 110 80 85 90
 E--> 112 GAT ACT GTT GAA GAG ACG GTG TTC AAT ACT TTC AAA TGT GTC ATG
 113 Asp Thr Val Glu Glu Thr Val Phe Asn Thr Phe Lys Cys Val Met
 114 95 100 105
 E--> 116 AAA AAC AAG CCA AAG TTC TCA CCA GTT GAT TGA ACCACCACGA
 117 Lys Asn Lys Pro Lys Phe Ser Pro Val Asp *
 118 110 115
 E--> 120 CTAGTAGATG GTTCAAATGG TGTGCTTTAC ATATAAAAT AAAGTGTTTC
 E--> 122 TGATGTAAAA AAAAAAAAAA AAAAAAAAAA AACTCGAGAG TATTCTAGAG
 E--> 124 CGGCCGCGGG CCCATCGTTT TCCACCC
 C--> 127 (2) INFORMATION FOR SEQ ID NO: 3
 128 (i) SEQUENCE CHARACTERISTICS:
 129 (A) LENGTH: 134 Amino Acids
 130 (B) TYPE: Amino Acid
 131 (C) STRANDEDNESS: single
 132 (D) TOPOLOGY: linear
 133 (ii) MOLECULE TYPE: Protein
 134 (iii) HYPOTHETICAL: no
 135 (iv) ANTI-SENSE: no
 136 (vi) ORIGINAL SOURCE:

delete if no sense, the valid heading

(36) is (ix) FEATURE?

46

↓

nos,

off

↓

P3

RAW SEQUENCE LISTING

DATE: 06/27/2001

PATENT APPLICATION: US/09/876,796

TIME: 15:33:37

Input Set : A:\SEQLIST.txt

Output Set: N:\CRF3\06272001\I876796.raw

```

137      (A) ORGANISM: Tenebrio molitor
C--> 138      (C) INDIVIDUAL ISOLATE: none
C--> 139      (G) CELL TYPE: fat body and whole organism
140      (vii) IMMEDIATE SOURCE:
141      (A) LIBRARY: cDNA
142      (B) CLONE: 13.17
143      (ix) FEATURES
145      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
147 Met Lys Leu Leu Cys Cys Leu Ile Ser Leu Ile Leu Leu Val Thr Val
148      -15      -10      -5
150 Gln Ala Leu Thr Glu Ala Gln Ile Glu Lys Leu Asn Lys Ile Ser Lys
E--> 151 1 5 5 10 10 15
153 Lys Cys Gln Asn Glu Ser Gly Val Ser Gln Glu Ile Ile Thr Lys Ala
154 15 20 25 30
156 Arg Asn Gly Asp Trp Glu Asp Asp Pro Lys Leu Lys Arg Gln Val Phe
157      35      40      45
159 Cys Val Ala Arg Asn Ala Gly Leu Ala Thr Glu Ser Gly Glu Val Val
160      50      55      60
162 Val Asp Val Leu Arg Glu Lys Val Arg Lys Val Thr Asp Asn Asp Glu
163      65      70      75
165 Glu Thr Glu Lys Ile Ile Asn Lys Cys Ala Val Lys Arg Asp Thr Val
166      80      85      90
168 Glu Glu Thr Val Phe Asn Thr Phe Lys Cys Val Met Lys Asn Lys Pro
169 95 100 105 110
E--> 171 Lys Phe Ser Pro Val Asp *
E--> 172      115
C--> 175 (2) INFORMATION FOR SEQ ID NO: 4
176      (i) SEQUENCE CHARACTERISTICS:
177      (A) LENGTH: 116 Amino Acids
178      (B) TYPE: Amino Acid
179      (C) STRANDEDNESS: single
180      (D) TOPOLOGY: linear
181      (ii) MOLECULE TYPE: Protein
182      (iii) HYPOTHETICAL: no
183      (iv) ANTI-SENSE: no
184      (vi) ORIGINAL SOURCE:
185      (A) ORGANISM: Tenebrio molitor
C--> 186      (C) INDIVIDUAL ISOLATE: none
C--> 187      (G) CELL TYPE: fat body and whole organism
188      (vii) IMMEDIATE SOURCE:
189      (A) LIBRARY: cDNA
190      (B) CLONE: 13.17
191      (ix) FEATURES
193      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
195 Leu Thr Glu Ala Gln Ile Glu Lys Leu Asn Lys Ile Ser Lys Lys Cys
196 1 5 10 15
198 Gln Asn Glu Ser Gly Val Ser Gln Glu Ile Ile Thr Lys Ala Arg Asn
199      20      25      30
201 Gly Asp Trp Glu Asp Asp Pro Lys Leu Lys Arg Gln Val Phe Cys Val

```

nos.
off

P. 4

RAW SEQUENCE LISTING

DATE: 06/27/2001

PATENT APPLICATION: US/09/876,796

TIME: 15:33:37

Input Set : A:\SEQLIST.txt

Output Set: N:\CRF3\06272001\I876796.raw

```

202          35          40          45
204 Ala Arg Asn Ala Gly Leu Ala Thr Glu Ser Gly Glu Val Val Val Asp
205          50          55          60
207 Val Leu Arg Glu Lys Val Arg Lys Val Thr Asp Asn Asp Glu Glu Thr
208 65          70          75          80
210 Glu Lys Ile Ile Asn Lys Cys Ala Val Lys Arg Asp Thr Val Glu Glu
211          85          90          95
213 Thr Val Phe Asn Thr Phe Lys Cys Val Met Lys Asn Lys Pro Lys Phe
214          100          105          110
E--> 216 Ser Pro Val Asp (*)
E--> 217          115
C--> 344 (2) INFORMATION FOR SEQ ID NO: 7
345      (i) SEQUENCE CHARACTERISTICS:
346          (A) LENGTH: 133 Amino Acids
347          (B) TYPE: Amino Acid
348          (C) STRANDEDNESS: single
349          (D) TOPOLOGY: linear
350      (ii) MOLECULE TYPE: Protein
351      (iii) HYPOTHETICAL: no
352      (iv) ANTI-SENSE: no
353      (vi) ORIGINAL SOURCE:
354          (A) ORGANISM: Tenebrio molitor
C--> 355      (C) INDIVIDUAL ISOLATE: none
C--> 356      (G) CELL TYPE: fat body and whole organism
357      (vii) IMMEDIATE SOURCE:
358          (A) LIBRARY: cDNA
359          (B) CLONE: 2.2, 2.3, and 7.5
360      (ix) FEATURES
362      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:
364 Met Lys Leu Leu Cys Phe Ala Phe Ala Ala Ile Val Ile Gly Ala
365          -15          -10          -5
367 Gln Ala Leu Thr Asp Glu Gln Ile Gln Lys Arg Asn Lys Ile Ser Lys
368          1          5          10
370 Glu Cys Gln Gln Val Ser Gly Val Ser Gln Glu Thr Ile Asp Lys Val
371 15          20          25          30
373 Arg Thr Gly Val Leu Val Asp Asp Pro Lys Met Lys Lys His Val Leu
374          35          40          45
376 Cys Phe Ser Lys Lys Thr Gly Val Ala Thr Glu Ala Gly Asp Thr Asn
377          50          55          60
379 Val Glu Val Leu Lys Ala Lys Leu Lys His Val Ala Ser Asp Glu Glu
380          65          70          75
382 Val Asp Lys Ile Val Gln Lys Cys Val Val Lys Lys Ala Thr Pro Glu
383          80          85          90
385 Glu Thr Ala Tyr Asp Thr Phe Lys Cys Ile Tyr Asp Ser Lys Pro Asp
386 95          100          105          110
E--> 388 Phe Ser Pro Ile Asp (*)
E--> 389          115
C--> 392 (2) INFORMATION FOR SEQ ID NO: 8
393      (i) SEQUENCE CHARACTERISTICS:

```

P S

RAW SEQUENCE LISTING

DATE: 06/27/2001

PATENT APPLICATION: US/09/876,796

TIME: 15:33:38

Input Set : A:\SEQLIST.txt

Output Set: N:\CRF3\06272001\I876796.raw

```

394      (A) LENGTH: 115 Amino Acids
395      (B) TYPE: Amino Acid
396      (C) STRANDEDNESS: single
397      (D) TOPOLOGY: linear
398      (ii) MOLECULE TYPE: Protein
399      (iii) HYPOTHETICAL: no
400      (iv) ANTI-SENSE: no
401      (vi) ORIGINAL SOURCE:
402          (A) ORGANISM: Tenebrio molitor
C--> 403      (C) INDIVIDUAL ISOLATE: none
C--> 404      (G) CELL TYPE: fat body and whole organism
405      (vii) IMMEDIATE SOURCE:
406          (A) LIBRARY: cDNA
407          (B) CLONE: 2.2, 2.3, and 7.5
408          (ix) FEATURES
410      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:
412 Leu Thr Asp Gln Gln Ile Gln Lys Arg Asn Lys Ile Ser Lys Glu Cys
413 1      5      10      15
415 Gln Gln Val Ser Gly Val Ser Gln Glu Thr Ile Asp Lys Val Arg Thr
416      20      25      30
418 Gly Val Leu Val Asp Asp Pro Lys Met Lys Lys His Val Leu Cys Phe
419      35      40      45
421 Ser Lys Lys Thr Gly Val Ala Thr Glu Ala Gly Asp Thr Asn Val Glu
422      50      55      60
424 Val Leu Lys Ala Lys Leu Lys His Val Ala Ser Asp Glu Glu Val Asp
425 65      70      75      80
427 Lys Ile Val Gln Lys Cys Val Val Lys Lys Ala Thr Pro Glu Glu Thr
428      85      90      95
430 Ala Tyr Asp Thr Phe Lys Cys Ile Tyr Asp Ser Lys Pro Asp Phe Ser
431      100     105     110
E--> 433 Pro Ile Asp *
E--> 434      115
C--> 500 (2) INFORMATION FOR SEQ ID NO: 10
501      (i) SEQUENCE CHARACTERISTICS:
502          (A) LENGTH: 133 Amino Acids
503          (B) TYPE: Amino Acid
504          (C) STRANDEDNESS: single
505          (D) TOPOLOGY: linear
506      (ii) MOLECULE TYPE: Protein
507      (iii) HYPOTHETICAL: no
508      (iv) ANTI-SENSE: no
509      (vi) ORIGINAL SOURCE:
510          (A) ORGANISM: Tenebrio molitor
C--> 511      (C) INDIVIDUAL ISOLATE: none
C--> 512      (G) CELL TYPE: fat body and whole organism
513      (vii) IMMEDIATE SOURCE:
514          (A) LIBRARY: cDNA
515          (B) CLONE: 3.4
516          (ix) FEATURES

```

P. 6

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/876,796

DATE: 06/27/2001

TIME: 15:33:38

Input Set : A:\SEQLIST.txt

Output Set: N:\CRF3\06272001\I876796.raw

```

518      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:
520 Met Lys Leu Leu Cys Phe Ala Phe Ala Ala Ile Val Ile Gly Ala
521      -15                      -10                      -5
523 Gln Ala Leu Thr Asp Glu Gln Ile Gln Lys Arg Asn Lys Ile Ser Lys
524      1                      5                      10
526 Glu Cys Gln Gln Val Ser Gly Val Ser Gln Glu Thr Ile Asp Lys Val
527 15                      20                      25                      30
529 Arg Thr Gly Val Leu Val Asp Asp Pro Lys Met Lys Lys His Val Leu
530      35                      40                      45
532 Cys Phe Ser Lys Lys Thr Gly Val Ala Thr Glu Ala Gly Asp Thr Asn
533      50                      55                      60
535 Val Glu Val Leu Lys Ala Lys Leu Lys His Val Ala Ser Asp Glu Glu
536      65                      70                      75
538 Val Asp Lys Ile Val Gln Lys Cys Val Val Lys Lys Ala Thr Pro Glu
539      80                      85                      90
541 Glu Thr Ala Tyr Asp Thr Phe Lys Val Ile Tyr Asp Ser Lys Pro Asp
542 95                      100                      105                      110
E--> 544 Phe Ser Pro Ile Asp *
E--> 545      115
C--> 548 (2) INFORMATION FOR SEQ ID NO: 11
549      (i) SEQUENCE CHARACTERISTICS:
550          (A) LENGTH: 115 Amino Acids
551          (B) TYPE: Amino Acid
552          (C) STRANDEDNESS: single
553          (D) TOPOLOGY: linear
554      (ii) MOLECULE TYPE: Protein
555      (iii) HYPOTHETICAL: no
556      (iv) ANTI-SENSE: no
557      (vi) ORIGINAL SOURCE:
558          (A) ORGANISM: Tenebrio molitor
C--> 559          (C) INDIVIDUAL ISOLATE: none
C--> 560          (G) CELL TYPE: fat body and whole organism
561      (vii) IMMEDIATE SOURCE:
562          (A) LIBRARY: cDNA
563          (B) CLONE: 3.4
564      (ix) FEATURES
566      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:
568 Leu Thr Asp Glu Gln Ile Gln Lys Arg Asn Lys Ile Ser Lys Glu Cys
569 1                      5                      10                      15
571 Gln Gln Val Ser Gly Val Ser Gln Glu Thr Ile Asp Lys Val Arg Thr
572      20                      25                      30
574 Gly Val Leu Val Asp Asp Pro Lys Met Lys Lys His Val Leu Cys Phe
575      35                      40                      45
577 Ser Lys Lys Thr Gly Val Ala Thr Glu Ala Gly Asp Thr Asn Val Glu
578      50                      55                      60
580 Val Leu Lys Ala Lys Leu Lys His Val Ala Ser Asp Glu Glu Val Asp
581 65                      70                      75                      80
583 Lys Ile Val Gln Lys Cys Val Val Lys Lys Ala Thr Pro Glu Glu Thr
584      85                      90                      95

```

RAW SEQUENCE LISTING

DATE: 06/27/2001

PATENT APPLICATION: US/09/876,796

TIME: 15:33:38

Input Set : A:\SEQLIST.txt

Output Set: N:\CRF3\06272001\I876796.raw

```

586 Ala Tyr Asp Thr Phe Lys Val Ile Tyr Asp Ser Lys Pro Asp Phe Ser
587                               100           105           110
E--> 589 Pro Ile Asp (*)
E--> 590                               115
C--> 656 (2) INFORMATION FOR SEQ ID NO: 13
657     (i) SEQUENCE CHARACTERISTICS:
658         (A) LENGTH: 133 Amino Acids
659         (B) TYPE: Amino Acid
660         (C) STRANDEDNESS: single
661         (D) TOPOLOGY: linear
662     (ii) MOLECULE TYPE: Protein
663     (iii) HYPOTHETICAL: no
664     (iv) ANTI-SENSE: no
665     (vi) ORIGINAL SOURCE:
666         (A) ORGANISM: Tenebrio molitor
C--> 667     (C) INDIVIDUAL ISOLATE: none
C--> 668     (G) CELL TYPE: fat body and whole organism
669     (vii) IMMEDIATE SOURCE:
670         (A) LIBRARY: cDNA
671         (B) CLONE: 3.9
672     (ix) FEATURES
674     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:
676 Met Lys Leu Leu Leu Cys Phe Ala Phe Ala Ala Ile Val Ile Gly Ala
677     -15           -10           -5
679 Gln Ala Leu Thr Asp Glu Gln Ile Gln Lys Arg Asn Lys Ile Ser Lys
680     1           5           10
682 Glu Cys Gln Gln Glu Ser Gly Val Ser Gln Glu Thr Ile Asp Lys Val
683 15           20           25           30
685 Arg Thr Gly Val Leu Val Asp Asp Pro Lys Met Lys Lys His Val Leu
686     35           40           45
688 Cys Phe Ser Lys Arg Thr Gly Val Ala Thr Glu Ala Gly Asp Thr Asn
689     50           55           60
E--> 691 Val Glu Val Leu Lys Ala Lys Leu Lys His Val Ala Ser Asp Glu Glu
E--> 692     65           70           75
694 Val Asp Lys Ile Val Gln Lys Cys Val Val Lys Lys Ala Thr Pro Glu
E--> 695     80           85           90
697 Glu Thr Ala Tyr Asp Thr Phe Lys Cys Ile Tyr Asp Ser Lys Pro Asp
E--> 698 95           100           105           110
E--> 700 Phe Ser Pro Ile Asp (*)
E--> 701     115
C--> 704 (2) INFORMATION FOR SEQ ID NO: 14
705     (i) SEQUENCE CHARACTERISTICS:
706         (A) LENGTH: 115 Amino Acids
707         (B) TYPE: Amino Acid
708         (C) STRANDEDNESS: single
709         (D) TOPOLOGY: linear
710     (ii) MOLECULE TYPE: Protein
711     (iii) HYPOTHETICAL: no
712     (iv) ANTI-SENSE: no

```

RAW SEQUENCE LISTING

DATE: 06/27/2001

PATENT APPLICATION: US/09/876,796

TIME: 15:33:38

Input Set : A:\SEQLIST.txt

Output Set: N:\CRF3\06272001\I876796.raw

```

713      (vi) ORIGINAL SOURCE:
714          (A) ORGANISM: Tenebrio molitor
C--> 715          (C) INDIVIDUAL ISOLATE: none
C--> 716          (G) CELL TYPE: fat body and whole organism
717      (vii) IMMEDIATE SOURCE:
718          (A) LIBRARY: cDNA
719          (B) CLONE: 3.9
720          (ix) FEATURES
722      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:
724 Leu Thr Asp Glu Gln Ile Gln Lys Arg Asn Lys Ile Ser Lys Glu Cys
725 1          5          10          15
727 Gln Gln Glu Ser Gly Val Ser Gln Glu Thr Ile Asp Lys Val Arg Thr
728          20          25          30
730 Gly Val Leu Val Asp Asp Pro Lys Met Lys Lys His Val Leu Cys Phe
731          35          40          45
733 Ser Lys Arg Thr Gly Val Ala Thr Glu Ala Gly Asp Thr Asn Val Glu
734          50          55          60
E--> 736 Val Leu Lys Ala Lys Leu Lys HisVal Ala Ser Asp Glu Glu Val Asp
E--> 737 65          70          75          80
739 Lys Ile Val Gln Lys Cys Val Val Lys Lys Ala Thr Pro Glu Glu Thr
E--> 740          85          90          95
742 Ala Tyr Asp Thr Phe Lys Cys Ile Tyr Asp Ser Lys Pro Asp Phe Ser
E--> 743          100          105          110
E--> 745 Pro Ile Asp *
E--> 746          115
C--> 749 (2) INFORMATION FOR SEQ ID NO: 15
750      (i) SEQUENCE CHARACTERISTICS:
751          (A) LENGTH: 481 base pairs
752          (B) TYPE: nucleic acid
753          (C) STRANDEDNESS: double
754          (D) TOPOLOGY: linear
755      (ii) MOLECULE TYPE: cDNA to mRNA
756      (iii) HYPOTHETICAL: no
757      (iv) ANTI-SENSE: no
758      (vi) ORIGINAL SOURCE:
759          (A) ORGANISM: Tenebrio molitor
C--> 760          (C) INDIVIDUAL ISOLATE: none
C--> 761          (G) CELL TYPE: fat body and whole organism
762      (vii) IMMEDIATE SOURCE:
763          (A) LIBRARY: cDNA
764          (B) CLONE: 7.5
765          (ix) FEATURES
767      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:
769 GGCACGAGCA AAA ATG AAA CTC CTC TTG TGC TTT GCG TTC GCC GCC      46
770          Met Lys Leu Leu Leu Cys Phe Ala Phe Ala Ala
771          -15          -10
E--> 773 ATC GTC ATC GGA GCT CAG GCT CTC ACC GAC GAA CAG ATA CAG AAA
774 Ile Val Ile Gly Ala Gln Ala Leu Thr Asp Glu Gln Ile Gln Lys
775          -5          1          5

```

(90) 91

RAW SEQUENCE LISTING

DATE: 06/27/2001

PATENT APPLICATION: US/09/876,796

TIME: 15:33:38

Input Set : A:\SEQLIST.txt

Output Set: N:\CRF3\06272001\I876796.raw

```

777 AGG AAC AAG ATC AGC AAA GAG TGC CAG CAG GTG TCC GGA GTG TCC 136
778 Arg Asn Lys Ile Ser Lys Glu Cys Gln Gln Val Ser Gly Val Ser
779 10 15 20
781 CAA GAG ACG ATC GAC AAA GTC CGC ACA GGT GTC TTG GTC GAC GAT 181
782 Gln Glu Thr Ile Asp Lys Val Arg Thr Gly Val Leu Val Asp Asp
783 25 30 35
785 CCC AAA ATG AAG AAG CAC GTC CTC TGC TTC TCG AAG AAA ACT GGA 226
786 Pro Lys Met Lys Lys His Val Leu Cys Phe Ser Lys Lys Thr Gly
787 40 45 50
789 GTG GCA ACC GAA GCC GGA GAC ACC AAT GTG GAG GTA CTC AAA GCC 271
790 Val Ala Thr Glu Ala Gly Asp Thr Asn Val Glu Val Leu Lys Ala
791 55 60 65
793 AAG CTG AAG CAT GTG GCC AGC GAC GAA GAG GTG GAC AAG ATC GTG 316
794 Lys Leu Lys His Val Ala Ser Asp Glu Glu Val Asp Lys Ile Val
795 70 75 80
797 CAG AAG TGC GTG GTC AAG AAG GCC ACA CCA GAG GAA ACG GCT TAT 361
798 Gln Lys Cys Val Val Lys Lys Ala Thr Pro Glu Glu Thr Ala Tyr
799 85 90 95
801 GAC ACC TTC AAG TGT ATT TAC GAC AGT AAA CCT GAT TTC TCT CCT 406
802 Asp Thr Phe Lys Cys Ile Tyr Asp Ser Lys Pro Asp Phe Ser Pro
803 100 105 110
805 ATT GAT TAA TTGTTTTGTA TTTGGCTGAA TTTTGACAAT AAAGGTACTA 455
806 Ile Asp *
807 115
809 TCGTTATGTA AAAAAAAAAA AAAAAA 481
C--> 888 (2) INFORMATION FOR SEQ ID NO: 17
889 (i) SEQUENCE CHARACTERISTICS: 173
890 (A) LENGTH: 169 Amino Acids
891 (B) TYPE: Amino Acid
892 (C) STRANDEDNESS: single
893 (D) TOPOLOGY: linear P.10
894 (ii) MOLECULE TYPE: Protein
895 (iii) HYPOTHETICAL: no
896 (iv) ANTI-SENSE: no
897 (vi) ORIGINAL SOURCE:
898 (A) ORGANISM: Tenebrio molitor
C--> 899 (C) INDIVIDUAL ISOLATE: none
C--> 900 (G) CELL TYPE: fat body and whole organism
901 (vii) IMMEDIATE SOURCE:
902 (A) LIBRARY: cDNA
903 (B) CLONE: 2.2
904 (ix) FEATURES-
906 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:
908 Met Gly Ser Ser His His His His His Ser Ser Gly Leu Val Pro .
909 -55 -50 -45
911 Arg Gly Ser His Met Ala Ser Met Thr Gly Gly Gln Gln Met Gly Arg
912 -40 -35 -30
914 Gly Ser Glu Phe Ala Arg Ala Lys Met Lys Leu Leu Leu Cys Phe Ala
915 -25 -20 -15

```

RAW SEQUENCE LISTING

DATE: 06/27/2001

PATENT APPLICATION: US/09/876,796

TIME: 15:33:38

Input Set : A:\SEQLIST.txt

Output Set: N:\CRF3\06272001\I876796.raw

```

917 Phe Ala Ala Ile Val Ile Gly Ala Gln Ala Leu Thr Asp Glu Gln Ile
918 -10 -5 1 5
920 Gln Lys Arg Asn Lys Ile Ser Lys Glu Cys Gln Gln Val Ser Gly Val
921 10 15 20
923 Ser Gln Glu Thr Ile Asp Lys Val Arg Thr Gly Val Leu Val Asp Asp
924 25 30 35
926 Pro Lys Met Lys Lys His Val Leu Cys Phe Ser Lys Lys Thr Gly Val
927 40 45 50
929 Ala Thr Glu Ala Gly Asp Thr Asn Val Glu Val Leu Lys Ala Lys Leu
930 55 60 65 70
932 Lys His Val Ala Ser Asp Glu Glu Val Asp Lys Ile Val Gln Lys Cys
933 75 80 85
935 Val Val Lys Lys Ala Thr Pro Glu Glu Thr Ala Tyr Asp Thr Phe Lys
936 90 95 100
E--> 938 Cys Ile Tyr Asp Ser Lys Pro Asp Phe Ser Pro Ile Asp
E--> 939 105 110 115
C--> 1009 (2) INFORMATION FOR SEQ ID NO: 19
1010 (i) SEQUENCE CHARACTERISTICS:
1011 (A) LENGTH: 149 Amino Acids
1012 (B) TYPE: Amino Acid
1013 (C) STRANDEDNESS: single
1014 (D) TOPOLOGY: linear
1015 (ii) MOLECULE TYPE: Protein
1016 (iii) HYPOTHETICAL: no
1017 (iv) ANTI-SENSE: no
1018 (vi) ORIGINAL SOURCE:
1019 (A) ORGANISM: Tenebrio molitor
C--> 1020 (C) INDIVIDUAL ISOLATE: none
C--> 1021 (G) CELL TYPE: fat body and whole organism
1022 (vii) IMMEDIATE SOURCE:
1023 (A) LIBRARY: cDNA
1024 (B) CLONE: 2.2
1025 (ix) FEATURES
1027 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:
1029 Met Gly Ser Ser His His His His His Ser Ser Gly Leu Val Pro
1030 -30 -25 -20
1032 Arg Gly Ser His Met Ala Ser Met Thr Gly Gly Gln Gln Met Gly Arg
1033 -15 -10 -5
1035 Gly Ser Leu Thr Asp Glu Gln Ile Gln Lys Arg Asn Lys Ile Ser Lys
1036 1 5 10
1038 Glu Cys Gln Gln Val Ser Gly Val Ser Gln Glu Thr Ile Asp Lys Val
1039 15 20 25 30
1041 Arg Thr Gly Val Leu Val Asp Asp Pro Lys Met Lys Lys His Val Leu
1042 35 40 45
1044 Cys Phe Ser Lys Lys Thr Gly Val Ala Thr Glu Ala Gly Asp Thr Asn
1045 50 55 60
1047 Val Glu Val Leu Lys Ala Lys Leu Lys His Val Ala Ser Asp Glu Glu
1048 65 70 75
1050 Val Asp Lys Ile Val Gln Lys Cys Val Val Lys Lys Ala Thr Pro Glu

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/876,796

DATE: 06/27/2001

TIME: 15:33:38

Input Set : A:\SEQLIST.txt

Output Set: N:\CRF3\06272001\I876796.raw

```

1051      80                      85                      90
1053 Glu Thr Ala Tyr Asp Thr Phe Lys Cys Ile Tyr Asp Ser Lys Pro Asp
1054 95                      100                      105                      110
E--> 1056 Phe Ser Pro Ile Asp *
E--> 1057                      115
C--> 1136 (2) INFORMATION FOR SEQ ID NO: 21
1137      (i) SEQUENCE CHARACTERISTICS:
1138          (A) LENGTH: 169 Amino Acids
1139          (B) TYPE: Amino Acid
1140          (C) STRANDEDNESS: single
1141          (D) TOPOLOGY: linear
1142      (ii) MOLECULE TYPE: Protein
1143      (iii) HYPOTHETICAL: no
1144      (iv) ANTI-SENSE: no
1145      (vi) ORIGINAL SOURCE:
1146          (A) ORGANISM: Tenebrio molitor
C--> 1147      (C) INDIVIDUAL ISOLATE: none
C--> 1148      (G) CELL TYPE: fat body and whole organism
1149      (vii) IMMEDIATE SOURCE:
1150          (A) LIBRARY: cDNA
1151          (B) CLONE: 2.3
1152      (ix) FEATURES
1154      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:
1156 Met Gly Ser Ser His His His His His Ser Ser Gly Leu Val Pro
1157      -55                      -50                      -45
1159 Arg Gly Ser His Met Ala Ser Met Thr Gly Gly Gln Gln Met Gly Arg
1160      -40                      -35                      -30
1162 Gly Ser Glu Phe Ala Arg Ala Lys Met Lys Leu Leu Leu Cys Phe Ala
1163      -25                      -20                      -15
1165 Phe Ala Ala Ile Val Ile Gly Ala Gln Ala Leu Thr Asp Glu Gln Ile
1166 -10                      -5                      1                      5
1168 Gln Lys Arg Asn Lys Ile Ser Lys Glu Cys Gln Gln Val Ser Gly Val
1169      10                      15                      20
1171 Ser Gln Glu Thr Ile Asp Lys Val Arg Thr Gly Val Leu Val Asp Asp
1172      25                      30                      35
1174 Pro Lys Met Lys Lys His Val Leu Cys Phe Ser Lys Lys Thr Gly Val
1175      40                      45                      50
1177 Ala Thr Glu Ala Gly Asp Thr Asn Val Glu Val Leu Lys Ala Lys Leu
1178 55                      60                      65                      70
1180 Lys His Val Ala Ser Asp Glu Glu Val Asp Lys Ile Val Gln Lys Cys
1181      75                      80                      85
1183 Val Val Lys Lys Ala Thr Pro Glu Glu Thr Ala Tyr Asp Thr Phe Lys
1184      90                      95                      100
E--> 1186 Cys Ile Tyr Asp Ser Lys Pro Asp Phe Ser Pro Ile Asp *
E--> 1187      105                      110                      115
C--> 1257 (2) INFORMATION FOR SEQ ID NO: 23
1258      (i) SEQUENCE CHARACTERISTICS:
1259          (A) LENGTH: 149 Amino Acids
1260          (B) TYPE: Amino Acid

```

RAW SEQUENCE LISTING

DATE: 06/27/2001

PATENT APPLICATION: US/09/876,796

TIME: 15:33:38

Input Set : A:\SEQLIST.txt

Output Set: N:\CRF3\06272001\I876796.raw

```

1261      (C) STRANDEDNESS: single
1262      (D) TOPOLOGY: linear
1263      (ii) MOLECULE TYPE: Protein
1264      (iii) HYPOTHETICAL: no
1265      (iv) ANTI-SENSE: no
1266      (vi) ORIGINAL SOURCE:
1267      (A) ORGANISM: Tenebrio molitor
C--> 1268      (C) INDIVIDUAL ISOLATE: none
C--> 1269      (G) CELL TYPE: fat body and whole organism
1270      (vii) IMMEDIATE SOURCE:
1271      (A) LIBRARY: cDNA
1272      (B) CLONE: 2.3
1273      (ix) FEATURES
1275      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:
1277 Met Gly Ser Ser His His His His Ser Ser Gly Leu Val Pro
1278      -30      -25      -20
1280 Arg Gly Ser His Met Ala Ser Met Thr Gly Gly Gln Gln Met Gly Arg
1281      -15      -10      -5
1283 Gly Ser Leu Thr Asp Glu Gln Ile Gln Lys Arg Asn Lys Ile Ser Lys
1284      1      5      10
1286 Glu Cys Gln Gln Val Ser Gly Val Ser Gln Glu Thr Ile Asp Lys Val
1287 15      20      25      30
1289 Arg Thr Gly Val Leu Val Asp Asp Pro Lys Met Lys Lys His Val Leu
1290      35      40      45
1292 Cys Phe Ser Lys Lys Thr Gly Val Ala Thr Glu Ala Gly Asp Thr Asn
1293      50      55      60
1295 Val Glu Val Leu Lys Ala Lys Leu Lys His Val Ala Ser Asp Glu Glu
1296      65      70      75
1298 Val Asp Lys Ile Val Gln Lys Cys Val Val Lys Lys Ala Thr Pro Glu
1299      80      85      90
1301 Glu Thr Ala Tyr Asp Thr Phe Lys Cys Ile Tyr Asp Ser Lys Pro Asp
1302 95      100      105      110
E--> 1304 Phe Ser Pro Ile Asp *
E--> 1305      115
C--> 1308 (2) INFORMATION FOR SEQ ID NO: 24
1309      (i) SEQUENCE CHARACTERISTICS:
1310      (A) LENGTH: 777 base pairs
1311      (B) TYPE: nucleic acid
1312      (C) STRANDEDNESS: double
1313      (D) TOPOLOGY: linear
1314      (ii) MOLECULE TYPE: cDNA to mRNA
1315      (iii) HYPOTHETICAL: no
1316      (iv) ANTI-SENSE: no
1317      (vi) ORIGINAL SOURCE:
1318      (A) ORGANISM: Tenebrio molitor
C--> 1319      (C) INDIVIDUAL ISOLATE: none
C--> 1320      (G) CELL TYPE: fat body and whole organism
1321      (vii) IMMEDIATE SOURCE:
1322      (A) LIBRARY: cDNA

```

DATE: 06/27/2001

TIME: 15:33:38

Input Set : A:\SEQLIST.txt

Output Set: N:\CRF3\06272001\I876796.raw

Line	Seq ID	Seq Description	Seq	Protein	Position
1323	(B) CLONE: 13.17				
1324		(ix) FEATURES			
1327	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 24:				
1329		TTGTTAGCGG ATGGAATTCC CTCGTAGGGG ATAATTTTGT TTACTTTAAG			50
1331		AAGGAGATAT ACC ATG GGC AGC AGC CAT CAT CAT CAT CAC AGC			96
1332		Met Gly Ser Ser His His His His His His Ser			
1333		-65 -60 -55			
1335		AGC GGC CTG GTG CCG CGC GGC AGC CAT ATG GCT AGC ATG ACT GGT			141
1336		Ser Gly Leu Val Pro Arg Gly Ser His Met Ala Ser Met Thr Gly			
1337		-50 -45 -40			
1339		GGA CAG CAA ATG GGT CGC GGA TCC GAA TTC TGG ATC CAA AGA ATT			186
1340		Gly Gln Gln Met Gly Arg Gly Ser Glu Phe Trp Ile Gln Arg Ile			
1341		-35 -30 -25			
1343		CGG CAC GAG ACT ACT AAG ATG AAG TTG CTC TGT TGT CTA ATC TCC			231
1344		Arg His Glu Thr Thr Lys Met Lys Leu Leu Cys Cys Leu Ile Ser			
1345		-20 -15 -10			
1348		CTC ATT CTG TTG GTC ACA GTT CAG GCC CTG ACC GAG GCA CAA ATT			276
1349		Leu Ile Leu Leu Val Thr Val Gln Ala Leu Thr Glu Ala Gln Ile			
1350		-5 1 5			
1352		GAG AAA CTG AAC AAG ATC AGC AAA AAA TGT CAA AAT GAA AGT GGA			321
1353		Glu Lys Leu Asn Lys Ile Ser Lys Lys Cys Gln Asn Glu Ser Gly			
1354		10 15 20			
1356		GTG TCG CAA GAG ATC ATA ACC AAA GCT CGC AAC GGT GAC TGG GAG			366
1357		Val Ser Gln Glu Ile Ile Thr Lys Ala Arg Asn Gly Asp Trp Glu			
1358		25 30 35			
1360		GAC GAT CCT AAA CTG AAA CGC CAA GTT TTT TGC GTG GCC AGG AAC			411
1361		Asp Asp Pro Lys Leu Lys Arg Gln Val Phe Cys Val Ala Arg Asn			
1362		40 45 50			
1364		GCC GGT CTG GCC ACG GAA TCG GGA GAG GTG GTG GTC GAC GTG TTG			456
1365		Ala Gly Leu Ala Thr Glu Ser Gly Glu Val Val Val Asp Val Leu			
1366		55 60 65			
1368		AGG GAG AAG GTG AGG AAG GTC ACT GAC AAC GAC GAA GAA ACT GAG			501
1369		Arg Glu Lys Val Arg Lys Val Thr Asp Asn Asp Glu Glu Thr Glu			
1370		70 75 80			
1372		AAA ATC ATC AAT AAG TGC GCC GTC AAG AGA GAT ACT GTT GAA GAG			546
1373		Lys Ile Ile Asn Lys Cys Ala Val Lys Arg Asp Thr Val Glu Glu			
1374		85 90 95			
1376		AAG GTG TTC AAT ACT TTC AAA TGT GTC ATG AAA AAC AAG CCA AAG			595
1377		Thr Val Phe Asn Thr Phe Lys Cys Val Met Lys Asn Lys Pro Lys			
1378		100 105 110			
1380		TTC TCA CCA GTT GAT TGA ACCACCACGA CTAGTAGATG GTTCAAATGG			643
1381		Phe Ser Pro Val Asp *			
1382		115			
1384		TGTGCTTTAC ATATAAAAAAT AAAGTGTTTC TGATGTAAAA AAAAAAAAAA			693
1386		AAAAAAAAAA AACTCGAGAG TATTCTAGAG CGGCCGCGGG CCCATCGTTT			743
1388		TCCACCCCTC GAGCACCACC ACCACCACCA CTGAGAT			777
1391	(2) INFORMATION FOR SEQ ID NO: 25				
1392	(i) SEQUENCE CHARACTERISTICS:				
1393	(A) LENGTH: 170 Amino Acids				

174 p. 14

595 591
643 nos. off
693

RAW SEQUENCE LISTING

DATE: 06/27/2001

PATENT APPLICATION: US/09/876,796

TIME: 15:33:38

Input Set : A:\SEQLIST.txt

Output Set: N:\CRF3\06272001\I876796.raw

```

1394      (B) TYPE: Amino Acid
1395      (C) STRANDEDNESS: single
1396      (D) TOPOLOGY: linear
1397      (ii) MOLECULE TYPE: Protein
1398      (iii) HYPOTHETICAL: no
1399      (iv) ANTI-SENSE: no
1400      (vi) ORIGINAL SOURCE:
1401          (A) ORGANISM: Tenebrio molitor
C--> 1402      (C) INDIVIDUAL ISOLATE: none
C--> 1403      (G) CELL TYPE: fat body and whole organism
1404      (vii) IMMEDIATE SOURCE:
1405          (A) LIBRARY: cDNA
1406          (B) CLONE: 13.17
1407          (ix) FEATURES
1409      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 25:
1411 Met Gly Ser Ser His His His His His Ser Ser Gly Leu Val Pro
1412      -55 -50 -45
1414 Arg Gly Ser His Met Ala Ser Met Thr Gly Gly Gln Gln Met Gly Arg
1415      -40 -35 -30
1417 Gly Ser Glu Phe Ala Arg Ala Lys Met Lys Leu Leu Cys Cys Leu Ile
1418      -25 -20 -15 -10
1420 Ser Leu Ile Leu Leu Val Thr Val Gln Ala Leu Thr Glu Ala Gln Ile
E--> 1421      -5 1 1 5 ← hos-off
1423 Glu Lys Leu Asn Lys Ile Ser Lys Lys Cys Gln Asn Glu Ser Gly Val
1424      10 15 20
1426 Ser Gln Glu Ile Ile Thr Lys Ala Arg Asn Gly Asp Trp Glu Asp Asp
1427      25 30 35
1429 Pro Lys Leu Lys Arg Gln Val Phe Cys Val Ala Arg Asn Ala Gly Leu
1430      40 45 50
1432 Ala Thr Glu Ser Gly Glu Val Val Val Asp Val Leu Arg Glu Lys Val
1433      55 60 65 70
1435 Arg Lys Val Thr Asp Asn Asp Glu Glu Thr Glu Lys Ile Ile Asn Lys
1436      75 80 85
1438 Cys Ala Val Lys Arg Asp Thr Val Glu Glu Thr Val Phe Asn Thr Phe
1439      90 95 100
E--> 1441 Lys Cys Val Met Lys Asn Lys Pro Lys Phe Ser Pro Val Asp (*)
E--> 1442      105 110 115
C--> 1512 (2) INFORMATION FOR SEQ ID NO: 27
1513      (i) SEQUENCE CHARACTERISTICS:
1514          (A) LENGTH: 149 Amino Acids
1515          (B) TYPE: Amino Acid
1516          (C) STRANDEDNESS: single
1517          (D) TOPOLOGY: linear
1518      (ii) MOLECULE TYPE: Protein
1519      (iii) HYPOTHETICAL: no
1520      (iv) ANTI-SENSE: no
1521      (vi) ORIGINAL SOURCE:
1522          (A) ORGANISM: Tenebrio molitor
C--> 1523      (C) INDIVIDUAL ISOLATE: none

```

RAW SEQUENCE LISTING

DATE: 06/27/2001

PATENT APPLICATION: US/09/876,796

TIME: 15:33:38

Input Set : A:\SEQLIST.txt

Output Set: N:\CRF3\06272001\I876796.raw

C--> 1524 (G) CELL TYPE: fat body and whole organism
 1525 (vii) IMMEDIATE SOURCE:
 1526 (A) LIBRARY: cDNA
 1527 (B) CLONE: 13.17
 1528 (ix) FEATURES
 1530 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 27:
 1532 Met Gly Ser Ser His His His His His His Ser Ser Gly Leu Val Pro
 1533 -30 -25 -20
 1535 Arg Gly Ser His Met Ala Ser Met Thr Gly Gly Gln Gln Met Gly Arg
 1536 -15 -10 -5
 1538 Gly Leu Thr Glu Ala Gln Ile Glu Lys Leu Asn Lys Ile Ser Lys Lys
 1539 1 5 10 15
 1541 Cys Gln Asn Glu Ser Gly Val Ser Gln Glu Ile Ile Thr Lys Ala Arg
 1542 20 25 30
 1544 Asn Gly Asp Trp Glu Asp Asp Pro Lys Leu Lys Arg Gln Val Phe Cys
 1545 35 40 45
 1547 Val Ala Arg Asn Ala Gly Leu Ala Thr Glu Ser Gly Glu Val Val Val
 1548 50 55 60
 1550 Asp Val Leu Arg Glu Lys Val Arg Lys Val Thr Asp Asn Asp Glu Glu
 1551 65 70 75
 1553 Thr Glu Lys Ile Ile Asn Lys Cys Ala Val Lys Arg Asp Thr Val Glu
 1554 80 85 90 95
 1556 Glu Thr Val Phe Asn Thr Phe Lys Cys Val Met Lys Asn Lys Pro Lys
 1557 100 105 110
 E--> 1559 Phe Ser Pro Val Asp * *delita*
 E--> 1560 115
 C--> 1639 (2) INFORMATION FOR SEQ ID NO: 29
 1640 (i) SEQUENCE CHARACTERISTICS:
 1641 (A) LENGTH: 173 Amino Acids
 1642 (B) TYPE: Amino Acid
 1643 (C) STRANDEDNESS: single
 1644 (D) TOPOLOGY: linear
 1645 (ii) MOLECULE TYPE: Protein
 1646 (iii) HYPOTHETICAL: no
 1647 (iv) ANTI-SENSE: no
 1648 (vi) ORIGINAL SOURCE:
 1649 (A) ORGANISM: Tenebrio molitor
 C--> 1650 (C) INDIVIDUAL ISOLATE: none
 C--> 1651 (G) CELL TYPE: fat body and whole organism
 1652 (vii) IMMEDIATE SOURCE:
 1653 (A) LIBRARY: cDNA
 1654 (B) CLONE: 3.4
 C--> 1655 (ix) FEATURE:
 1656 (D) OTHER INFORMATION: Precursor protein with His-tag
 1657 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 29:
 1659 Met Gly Ser Ser His His His His His Ser Ser Gly Leu Val Pro
 1660 -55 -50 -45
 1662 Arg Gly Ser His Met Ala Ser Met Thr Gly Gly Gln Gln Met Gly Arg
 1663 -40 -35 -30

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/876,796

DATE: 06/27/2001

TIME: 15:33:38

Input Set : A:\SEQLIST.txt

Output Set: N:\CRF3\06272001\I876796.raw

```

1665 Gly Ser Glu Phe Ala Arg Ala Lys Met Lys Leu Leu Leu Cys Phe Ala
1666      -25                      -20                      -15
1668 Phe Ala Ala Ile Val Ile Gly Ala Gln Ala Leu Thr Asp Glu Gln Ile
1669 -10                      -5                      1                      5
1671 Gln Lys Arg Asn Lys Ile Ser Lys Glu Cys Gln Gln Val Ser Gly Val
1672      10                      15                      20
1674 Ser Gln Glu Thr Ile Asp Lys Val Arg Thr Gly Val Leu Val Asp Asp
1675      25                      30                      35
1677 Pro Lys Met Lys Lys His Val Leu Cys Phe Ser Lys Lys Thr Gly Val
1678      40                      45                      50
1680 Ala Thr Glu Ala Gly Asp Thr Asn Val Glu Val Leu Lys Ala Lys Leu
1681 55                      60                      65                      70
1683 Lys His Val Ala Ser Asp Glu Glu Val Asp Lys Ile Val Gln Lys Cys
1684      75                      80                      85
1686 Val Val Lys Lys Ala Thr Pro Glu Glu Thr Ala Tyr Asp Thr Phe Lys
1687      90                      95                      100
E--> 1689 Val Ile Tyr Asp Ser Lys Pro Asp Phe Ser Pro Ile Asp (*)
E--> 1690      105                      110                      115
C--> 1761 (2) INFORMATION FOR SEQ ID NO: 31
1762      (i) SEQUENCE CHARACTERISTICS:
1763          (A) LENGTH: 149 Amino Acids
1764          (B) TYPE: Amino Acid
1765          (C) STRANDEDNESS: single
1766          (D) TOPOLOGY: linear
1767      (ii) MOLECULE TYPE: Protein
1768      (iii) HYPOTHETICAL: no
1769      (iv) ANTI-SENSE: no
1770      (vi) ORIGINAL SOURCE:
1771          (A) ORGANISM: Tenebrio molitor
C--> 1772      (C) INDIVIDUAL ISOLATE: none
C--> 1773      (G) CELL TYPE: fat body and whole organism
1774      (vii) IMMEDIATE SOURCE:
1775          (A) LIBRARY: cDNA
1776          (B) CLONE: 3.4
C--> 1777      (ix) FEATURE:
1778          (D) OTHER INFORMATION: Mature Protein with His-tag
1779      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 31:
1781 Met Gly Ser Ser His His His His His Ser Ser Gly Leu Val Pro
1782      -30                      -25                      -20
1784 Arg Gly Ser His Met Ala Ser Met Thr Gly Gly Gln Gln Met Gly Arg
1785      -15                      -10                      -5
1787 Gly Ser Leu Thr Asp Glu Gln Ile Gln Lys Arg Asn Lys Ile Ser Lys
1788      1                      5                      10
1790 Glu Cys Gln Gln Val Ser Gly Val Ser Gln Glu Thr Ile Asp Lys Val
1791 15                      20                      25                      30
1793 Arg Thr Gly Val Leu Val Asp Asp Pro Lys Met Lys Lys His Val Leu
1794      35                      40                      45
1796 Cys Phe Ser Lys Lys Thr Gly Val Ala Thr Glu Ala Gly Asp Thr Asn
1797      50                      55                      60

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/876,796

DATE: 06/27/2001

TIME: 15:33:38

Input Set : A:\SEQLIST.txt

Output Set: N:\CRF3\06272001\I876796.raw

```

1799 Val Glu Val Leu Lys Ala Lys Leu Lys His Val Ala Ser Asp Glu Glu
1800      65      70      75
1802 Val Asp Lys Ile Val Gln Lys Cys Val Val Lys Lys Ala Thr Pro Glu
1803      80      85      90
1805 Glu Thr Ala Tyr Asp Thr Phe Lys Val Ile Tyr Asp Ser Lys Pro Asp
1806 95      100      105      110
E--> 1808 Phe Ser Pro Ile Asp *
E--> 1809      115
C--> 1888 (2) INFORMATION FOR SEQ ID NO: 33
1889      (i) SEQUENCE CHARACTERISTICS:
1890          (A) LENGTH: 173 Amino Acids
1891          (B) TYPE: Amino Acid
1892          (C) STRANDEDNESS: single
1893          (D) TOPOLOGY: linear
1894      (ii) MOLECULE TYPE: Protein
1895      (iii) HYPOTHETICAL: no
1896      (iv) ANTI-SENSE: no
1897      (vi) ORIGINAL SOURCE:
1898          (A) ORGANISM: Tenebrio molitor
C--> 1899      (C) INDIVIDUAL ISOLATE: none
C--> 1900      (G) CELL TYPE: fat body and whole organism
1901      (vii) IMMEDIATE SOURCE:
1902          (A) LIBRARY: cDNA
1903          (B) CLONE: 3.9
C--> 1904      (ix) FEATURE:
1905          (D) OTHER INFORMATION: Precursor Protein with His-tag
1906      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 33:
1908 Met Gly Ser Ser His His His His His His Ser Ser Gly Leu Val Pro
1909      -55      -50      -45
1911 Arg Gly Ser His Met Ala Ser Met Thr Gly Gly Gln Gln Met Gly Arg
1912      -40      -35      -30
1914 Gly Ser Glu Phe Ala Arg Ala Lys Met Lys Leu Leu Leu Cys Phe Ala
1915      -25      -20      -15
1917 Phe Ala Ala Ile Val Ile Gly Ala Gln Ala Leu Thr Asp Glu Gln Ile
1918 -10      -5      1      5
1920 Gln Lys Arg Asn Lys Ile Ser Lys Glu Cys Gln Gln Glu Ser Gly Val
1921      10      15      20
1923 Ser Gln Glu Thr Ile Asp Lys Val Arg Thr Gly Val Leu Val Asp Asp
1924      25      30      35
1926 Pro Lys Met Lys Lys His Val Leu Cys Phe Ser Lys Arg Thr Gly Val
1927      40      45      50
1929 Ala Thr Glu Ala Gly Asp Thr Asn Val Glu Val Leu Lys Ala Lys Leu
1930 55      60      65      70
E--> 1932 Lys HisVal Ala Ser Asp Glu Glu Val Asp Lys Ile Val Gln Lys Cys
E--> 1933      75      80      85
1935 Val Val Lys Lys Ala Thr Pro Glu Glu Thr Ala Tyr Asp Thr Phe Lys
E--> 1936      90      95
E--> 1938 Cys Ile Tyr Asp Ser Lys Pro Asp Phe Ser Pro Ile Asp *
E--> 1939      105      110      115

```

space

delete

RAW SEQUENCE LISTING

DATE: 06/27/2001

PATENT APPLICATION: US/09/876,796

TIME: 15:33:38

Input Set : A:\SEQLIST.txt

Output Set: N:\CRF3\06272001\I876796.raw

C--> 2009 (2) INFORMATION FOR SEQ ID NO: 35

2010 (i) SEQUENCE CHARACTERISTICS:

2011 (A) LENGTH: 149 Amino Acids

2012 (B) TYPE: Amino Acid

2013 (C) STRANDEDNESS: single

2014 (D) TOPOLOGY: linear

2015 (ii) MOLECULE TYPE: Protein

2016 (iii) HYPOTHETICAL: no

2017 (iv) ANTI-SENSE: no

2018 (vi) ORIGINAL SOURCE:

2019 (A) ORGANISM: Tenebrio molitor

C--> 2020 (C) INDIVIDUAL ISOLATE: none

C--> 2021 (G) CELL TYPE: fat body and whole organism

2022 (vii) IMMEDIATE SOURCE:

2023 (A) LIBRARY: cDNA

2024 (B) CLONE: 3.9

C--> 2025 (ix) FEATURE:

2026 (D) OTHER INFORMATION: Mature Protein with His-tag

2027 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 35:

2029 Met Gly Ser Ser His His His His His Ser Ser Gly Leu Val Pro

2030 -30 -25 -20

2032 Arg Gly Ser His Met Ala Ser Met Thr Gly Gly Gln Gln Met Gly Arg

2033 -15 -10 -5

2035 Gly Ser Leu Thr Asp Glu Gln Ile Gln Lys Arg Asn Lys Ile Ser Lys

2036 1 5 10

2038 Glu Cys Gln Gln Glu Ser Gly Val Ser Gln Glu Thr Ile Asp Lys Val

2039 15 20 25 30

2041 Arg Thr Gly Val Leu Val Asp Asp Pro Lys Met Lys Lys His Val Leu

2042 35 40 45

2044 Cys Phe Ser Lys Arg Thr Gly Val Ala Thr Glu Ala Gly Asp Thr Asn

2045 50 55 *space* 60

E--> 2047 Val Glu Val Leu Lys Ala Lys Leu Lys HisVal Ala Ser Asp Glu Glu

E--> 2048 65 70 75

2050 Val Asp Lys Ile Val Gln Lys Cys Val Val Lys Lys Ala Thr Pro Glu

E--> 2051 80 85 90

2053 Glu Thr Ala Tyr Asp Thr Phe Lys Cys Ile Tyr Asp Ser Lys Pro Asp

E--> 2054 95 100 105 110

E--> 2056 Phe Ser Pro Ile Asp *

E--> 2057 115

C--> 2139 (2) INFORMATION FOR SEQ ID NO: 37

2140 (i) SEQUENCE CHARACTERISTICS:

2141 (A) LENGTH: 173 Amino Acids

2142 (B) TYPE: Amino Acid

2143 (C) STRANDEDNESS: single

2144 (D) TOPOLOGY: linear

2145 (ii) MOLECULE TYPE: Protein

2146 (iii) HYPOTHETICAL: no

2147 (iv) ANTI-SENSE: no

2148 (vi) ORIGINAL SOURCE:

RAW SEQUENCE LISTING

DATE: 06/27/2001

PATENT APPLICATION: US/09/876,796

TIME: 15:33:38

Input Set : A:\SEQLIST.txt

Output Set: N:\CRF3\06272001\I876796.raw

2149 (A) ORGANISM: Tenebrio molitor
C--> 2150 (C) INDIVIDUAL ISOLATE: none
C--> 2151 (G) CELL TYPE: fat body and whole organism
2152 (vii) IMMEDIATE SOURCE:
2153 (A) LIBRARY: cDNA
2154 (B) CLONE: 7.5
C--> 2155 (ix) FEATURE:
2156 (D) OTHER INFORMATION: Precursor Protein with His-tag
2157 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 37
2159 Met Gly Ser Ser His His His His His Ser Ser Gly Leu Val Pro
2160 -55 -50 -45
2162 Arg Gly Ser His Met Ala Ser Met Thr Gly Gly Gln Gln Met Gly Arg
2163 -40 -35 -30
2165 Gly Ser Glu Phe Ala Arg Ala Lys Met Lys Leu Leu Cys Phe Ala
2166 -25 -20 -15
2168 Phe Ala Ala Ile Val Ile Gly Ala Gln Ala Leu Thr Asp Glu Gln Ile
2169 -10 -5 1 5
2171 Gln Lys Arg Asn Lys Ile Ser Lys Glu Cys Gln Gln Val Ser Gly Val
2172 10 15 20
2174 Ser Gln Glu Thr Ile Asp Lys Val Arg Thr Gly Val Leu Val Asp Asp
2175 25 30 35
2177 Pro Lys Met Lys Lys His Val Leu Cys Phe Ser Lys Lys Thr Gly Val
2178 40 45 50
2180 Ala Thr Glu Ala Gly Asp Thr Asn Val Glu Val Leu Lys Ala Lys Leu
2181 55 60 65 70
2183 Lys His Val Ala Ser Asp Glu Glu Val Asp Lys Ile Val Gln Lys Cys
2184 75 80 85
2186 Val Val Lys Lys Ala Thr Pro Glu Glu Thr Ala Tyr Asp Thr Phe Lys
2187 90 95 100
E--> 2189 Cys Ile Tyr Asp Ser Lys Pro Asp Phe Ser Pro Ile Asp *
E--> 2190 105 110 115
C--> 2260 (2) INFORMATION FOR SEQ ID NO: 39
2261 (i) SEQUENCE CHARACTERISTICS:
2262 (A) LENGTH: 149 Amino Acids
2263 (B) TYPE: Amino Acid
2264 (C) STRANDEDNESS: single
2265 (D) TOPOLOGY: linear
2266 (ii) MOLECULE TYPE: Protein
2267 (iii) HYPOTHETICAL: no
2268 (iv) ANTI-SENSE: no
2269 (vi) ORIGINAL SOURCE:
2270 (A) ORGANISM: Tenebrio molitor
C--> 2271 (C) INDIVIDUAL ISOLATE: none
C--> 2272 (G) CELL TYPE: fat body and whole organism
2273 (vii) IMMEDIATE SOURCE:
2274 (A) LIBRARY: cDNA
2275 (B) CLONE: 7.5
C--> 2276 (ix) FEATURE:
2277 (D) OTHER INFORMATION: Mature protein with His-tag

P. 20

RAW SEQUENCE LISTING

DATE: 06/27/2001

PATENT APPLICATION: US/09/876,796

TIME: 15:33:38

Input Set : A:\SEQLIST.txt

Output Set: N:\CRF3\06272001\I876796.raw

2278 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 39

2280 Met Gly Ser Ser His His His His Ser Ser Gly Leu Val Pro

2281 -30 -25 -20

2283 Arg Gly Ser His Met Ala Ser Met Thr Gly Gly Gln Gln Met Gly Arg

2284 -15 -10 -5

2286 Gly Ser Leu Thr Asp Glu Gln Ile Gln Lys Arg Asn Lys Ile Ser Lys

2287 1 5 10

2289 Glu Cys Gln Gln Val Ser Gly Val Ser Gln Glu Thr Ile Asp Lys Val

2290 15 20 25 30

2292 Arg Thr Gly Val Leu Val Asp Asp Pro Lys Met Lys Lys His Val Leu

2293 35 40 45

2295 Cys Phe Ser Lys Lys Thr Gly Val Ala Thr Glu Ala Gly Asp Thr Asn

2296 50 55 60

2298 Val Glu Val Leu Lys Ala Lys Leu Lys His Val Ala Ser Asp Glu Glu

2299 65 70 75

2301 Val Asp Lys Ile Val Gln Lys Cys Val Val Lys Lys Ala Thr Pro Glu

2302 80 85 90

2304 Glu Thr Ala Tyr Asp Thr Phe Lys Cys Ile Tyr Asp Ser Lys Pro Asp

2305 95 100 105 110

E--> 2307 Phe Ser Pro Ile Asp (*)

E--> 2308 115

C--> 2468 (2) INFORMATION FOR SEQ ID NO: 45

2469 (i) SEQUENCE CHARACTERISTICS: 484

2470 (A) LENGTH: 481 base pairs

2471 (B) TYPE: nucleic acid

2472 (C) STRANDEDNESS: double

2473 (D) TOPOLOGY: linear

2474 (ii) MOLECULE TYPE: cDNA to mRNA

2475 (iii) HYPOTHETICAL: no

2476 (iv) ANTI-SENSE: no

2477 (vi) ORIGINAL SOURCE:

2478 (A) ORGANISM: Tenebrio molitor

C--> 2479 (C) INDIVIDUAL ISOLATE: none

C--> 2480 (G) CELL TYPE: fat body and whole organism

2481 (vii) IMMEDIATE SOURCE:

2482 (A) LIBRARY: cDNA

2483 (B) CLONE: 2.2

C--> 2484 (ix) FEATURE: Consensus

2485 (D) OTHER INFORMATION: Consensus of Seq ID #44 with Tm 13.17

2486 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 45:

2488 GGCANRNNNN AAR ATG AAR YTN CTC TNN TGY YTN RYN TYC NYC RYY 46

2489 Met Lys Leu Leu Leu Cys Phe Ala Phe Ala Ala

2490 -15 -10

2492 NTN NTN RTC RNA GYT CAG GCY CTN ACC GAN GNA CAR ATN NAG AAA 91

2493 Ile Val Ile Gly Ala Gln Ala Leu Thr Asp Glu Gln Ile Gln Lys

2494 -5 1 5

2496 NNG AAC AAG ATC AGC AAA RAR TGY CAR NAN GNR NNY GGA GTG TCN 136

2497 Arg Asn Lys Ile Ser Lys Glu Cys Gln Gln Val Ser Gly Val Ser

2498 10 15 20

p.21

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/876,796

DATE: 06/27/2001

TIME: 15:33:38

Input Set : A:\SEQLIST.txt

Output Set: N:\CRF3\06272001\I876796.raw

```

2500 CAA GAG AYN ATN RNC AAA GYY CGC ANN GGT GNC TNG GNN GAY GAT      181
2501 Gln Glu Thr Ile Asp Lys Val Arg Thr Gly Val Leu Val Asp Asp
2502      25      30      35
2504 CCY AAA NTG AAR NRN CAN GTY YTY TGC NTN NCN ARG ARN RCY GGN      226
2505 Pro Lys Met Lys Lys His Val Leu Cys Phe Ser Lys Lys Thr Gly
2506      40      45      50
2508 NTG GCN ACN GAA NCN GGA GAN RYN RNN GTN GAN GTR YTN ARR GNN      271
2509 Val Ala Thr Glu Ala Gly Asp Thr Asn Val Glu Val Leu Lys Ala
2510      55      60      65
2512 AAG NTG ARG NAN GTN RCY RRC AAC GAC GAA GAR RYN GAN AAR ATC      316
2513 Lys Leu Lys His Val Ala Ser Asn Asp Glu Glu Val Asp Lys Ile
2514      70      75      80
2516 RTN NAN AAG TGC GYN GTC AAG ARR GNY ACN NYN GAR GAR ACG GYN      361
2517 Val Gln Lys Cys Val Val Lys Lys Ala Thr Pro Glu Glu Thr Ala
2518      85      90      95
2520 TNY RAY ACY TTC AAR NNT RTY NNN RAN ARY AAR CCN RAN TTC TCN      406
2521 Tyr Asp Thr Phe Lys Cys Ile Tyr Asp Ser Lys Pro Asp Phe Ser
2522      100      105      110
2524 CCN RTT GAT TRA NYNNYYNNNA YTNNGNNNRNR NTTYRANAAT AAAGNNNNNTN      458
2525 Pro Ile Asp *
2526      115
E--> 2528 TNRTNNNRNA AAAAAAAAAA AAAAAA
C--> 2533 (2) INFORMATION FOR SEQ ID NO: 46
2534      (i) SEQUENCE CHARACTERISTICS: 484
2535          (A) LENGTH: 481 base pairs
2536          (B) TYPE: nucleic acid
2537          (C) STRANDEDNESS: double
2538          (D) TOPOLOGY: linear
2539      (ii) MOLECULE TYPE: cDNA to mRNA
2540      (iii) HYPOTHETICAL: no
2541      (iv) ANTI-SENSE: no
2542          + delete
2543      (vi) ORIGINAL SOURCE:
2544          (A) ORGANISM: Tenebrio molitor
C--> 2545      (C) INDIVIDUAL ISOLATE: none
C--> 2546      (G) CELL TYPE: fat body and whole organism
2547      (vii) IMMEDIATE SOURCE:
2548          (A) LIBRARY: cDNA
2549          (B) CLONE: 2.2
C--> 2550      (ix) FEATURE:
2551          (D) OTHER INFORMATION: Consensus of Seq ID #45 with B1/B2
2552      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 46:
2554 GGCANRNNNN AAR ATG AAR YTN CTC TNN TGY YTN RYN TYY NYC RYY      46
2555      Met Lys Leu Leu Leu Cys Phe Ala Phe Ala Ala
2556      -15      -10
2558 NTN NTN RTC NNA GYT CAG GCY NTN ACY NAN GNA NAN NTN NAG NNA      91
2559 Ile Val Ile Gly Ala Gln Ala Leu Thr Asp Glu Gln Ile Gln Lys
2560      -5      1      5
2562 NNG NNC NAR AYC AGC RNA RAR TGY NAR NNN GNR NNY GGA GTG TCN      136

```

484

P.22

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/876,796

DATE: 06/27/2001

TIME: 15:33:38

Input Set : A:\SEQLIST.txt

Output Set: N:\CRF3\06272001\I876796.raw

```

2563 Arg Asn Lys Ile Ser Lys Glu Cys Gln Gln Val Ser Gly Val Ser
2564      10      15      20
2566 NAA GAN RYN ATN RNN ARA GYY CGC ANN GGT GNC TNG GNN GAY GAY      181
2567 Gln Glu Thr Ile Asp Lys Val Arg Thr Gly Val Leu Val Asp Asp
2568      25      30      35
2570 CCY AAA NTG AAR NNN CAN NTY YTY TGC NTN NYN ARG RNN NYI GRN      226
2571 Pro Lys Met Lys Lys His Val Leu Cys Phe Ser Lys Lys Thr Gly
2572      40      45      50
2574 NTR GYN RCN GAA NCN GGA GAN RYN RNN GYN GAN RYR YTN ARR GNN      271
2575 Val Ala Thr Glu Ala Gly Asp Thr Asn Val Glu Val Leu Lys Ala
2576      55      60      65
2578 AAG NTG ANG NRN NNN NNN RNN ANN RNN RAR RAR RYN RRN ARR NTN      316
2579 Lys Leu Lys His Val Ala Ser Asn Asp Glu Glu Val Asp Lys Ile
2580      70      75      80
2582 NYN NRN ARN NNN NNN NNN NNG ARN RNN NYN NNN RAR RNR NNN NNN      361
2583 Val Gln Lys Cys Val Val Lys Lys Ala Thr Pro Glu Glu Thr Ala
2584      85      90      95
2586 TNN RAN NYN YYN AAN NNN NNY NNN RRN ANN ARN CCN RNN TYY TYN      406
2587 Tyr Asp Thr Phe Lys Cys Ile Tyr Asp Ser Lys Pro Asp Phe Ser
2588     100     105     110
2590 CNN RYT RNT TRN NYNNNNNNNN YNNGNNNRNR NTTYRANAAT AAAGNNNYTN      458
2591 Pro Ile Asp *
2592     115

```

E--> 2594 TNRTNNNRNA AAAAAAAAAA AAAAAA

484

C--> 2599 (2) INFORMATION FOR SEQ ID NO: 47

2600 (i) SEQUENCE CHARACTERISTICS: 484

2601 (A) LENGTH: 481 base pairs

2602 (B) TYPE: nucleic acid

2603 (C) STRANDEDNESS: double

2604 (D) TOPOLOGY: linear P.23

2605 (ii) MOLECULE TYPE: cDNA to mRNA

2606 (iii) HYPOTHETICAL: no

2607 (iv) ANTI-SENSE: no

2608 (vi) ORIGINAL SOURCE:

2609 (A) ORGANISM: Tenebrio molitor

C--> 2610 (C) INDIVIDUAL ISOLATE: none

C--> 2611 (G) CELL TYPE: fat body and whole organism

2612 (vii) IMMEDIATE SOURCE:

2613 (A) LIBRARY: cDNA

2614 (B) CLONE: 2.2

C--> 2615 (ix) FEATURE: 5

2616 (D) OTHER INFORMATION: Consensus of Seq. ID #46 with AFP-3

2617 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 47:

2620 GGCNNRNNNN AAR ATG AAR YTN CTC YNN TGY YTN RYN YYY NYI RYY 46

2621 Met Lys Leu Leu Leu Cys Phe Ala Phe Ala Ala

2622 -15 -10

2624 NTN NTN RYC NNR RYY YAN GCY NTN ACY NAN RNA NNN NNN NAG NNR 91

2625 Ile Val Ile Gly Ala Gln Ala Leu Thr Asp Glu Gln Ile Gln Lys

2626 -5 1 5

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/876,796

DATE: 06/27/2001

TIME: 15:33:38

Input Set : A:\SEQLIST.txt

Output Set: N:\CRF3\06272001\I876796.raw

```

2628 NNG NNY NAR NNC AGC RNN RNN TGY NAR NNN GNR NNY GGA GTR TCN      136
2629 Arg Asn Lys Ile Ser Lys Glu Cys Gln Gln Glu Ser Gly Val Ser
2630      10      15      20
2632 NAA GAN NYN NTN RNN ARR GYY CGC ANN NGT GNN NNR GNN GAY GAY      181
2633 Gln Glu Thr Ile Asp Lys Val Arg Thr Gly Val Leu Val Asp Asp
2634      25      30      35
2636 CCY AAA NTG AAR NNN CAN NYI YTY TGC NTN NYN ARG RNN NYI GRN      226
2637 Pro Lys Met Lys Lys His Val Leu Cys Phe Ser Lys Lys Thr Gly
2638      40      45      50
2640 NTN RYN RNN GNN NNN GGN GAN NYN NNN NYN GAN NNN NTN ARR RNN      271
2641 Val Ala Thr Glu Ala Gly Asp Thr Asn Val Glu Val Leu Lys Ala
2642      55      60      65
2644 AAR NTN ANG NRN NNN NNN RNN RNN NNN RAR RAR RYN RRN RRN NTN      316
2645 Lys Leu Lys His Val Ala Ser Asn Asp Glu Glu Val Asp Lys Ile
2646      70      75      80
2648 NYN NNN ARN NNN NNN NNN NNG ARN RNN NYN NNN NAR NNN NNN NNN      361
2649 Val Gln Lys Cys Val Val Lys Lys Ala Thr Pro Glu Glu Thr Ala
2650      85      90      95
2652 NNN RAN NYN YYN AAN NNN NNY NNN RRN ANN ARN YCN NNN TNN NNN      406
2653 Tyr Asp Thr Phe Lys Cys Ile Tyr Asp Ser Lys Pro Asp Phe Ser
2654      100      105      110
2656 CNN NYN RNN TRN NNNNNNNNNN YNNRNNNNNN NNNNNNNAAT AAANNNNNNN      458
2657 Pro Ile Asp *
2658      115
E--> 2660 NNNNNNNNNNA AAAAAAAAAA AAAAAA
C--> 2664 (2) INFORMATION FOR SEQ ID NO: 48
2665      (i) SEQUENCE CHARACTERISTICS:
2666          (A) LENGTH: 133 Amino Acids
2667          (B) TYPE: Amino Acid
2668          (C) STRANDEDNESS: single
2669          (D) TOPOLOGY: linear
2670      (ii) MOLECULE TYPE: Protein
2671      (iii) HYPOTHETICAL: no
2672      (iv) ANTI-SENSE: no
2673      (vi) ORIGINAL SOURCE:
2674          (A) ORGANISM: Tenebrio molitor
C--> 2675      (C) INDIVIDUAL ISOLATE: none
C--> 2676      (G) CELL TYPE: fat body and whole organism
2677      (vii) IMMEDIATE SOURCE:
2678          (A) LIBRARY: cDNA
2679          (B) CLONE:
C--> 2680      (ix) FEATURE:
2681          (D) OTHER INFORMATION: General Consensus of Clones,
2682 B1, B2 and AFP-3
2684      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 48:
2686 Met Lys Leu Leu Cys Phe Ala Phe Ala Ala Ile Val Ile Gly Ala
E--> 2687      Cys      Leu Ile Ser Leu Ile Leu Leu Val Thr Val
E--> 2688      Thr Leu Val      Ala Ala Thr
E--> 2689      Val

```

484

(See 1.822 of
Sequence Rules.)

invalid grouping
of amino acids

General S

what is this?

RAW SEQUENCE LISTING

DATE: 06/27/2001

PATENT APPLICATION: US/09/876,796

TIME: 15:33:38

Input Set : A:\SEQLIST.txt

Output Set: N:\CRF3\06272001\I876796.raw

```

E--> 2690          -15          -10          -5
      2692 Gln Ala Leu Thr Asp Glu Gln Ile Gln Lys Arg Asn Lys Ile Ser Lys
E--> 2693 Tyr      Ile      Glu Ala Asp Leu Glu Leu Leu Arg Gln Thr Ala
E--> 2694          Thr Pro Arg          Lys      His      Asp
E--> 2695          1          5          10
      2697 Glu Cys Gln Gln Glu Ser Gly Val Ser Gln Glu Thr Ile Asp Lys Val
E--> 2698 Lys      Lys Asn Val      Ala      Glu Asp Ile Leu Thr Arg Ala
E--> 2699 Ala      Thr          Ala Val      Lys
E--> 2700          Ala          Ser      Asn
E--> 2701 15          20          25          30
      2703 Arg Thr Gly Val Leu Val Asp Asp Pro Lys Met Lys Lys His Val Leu
E--> 2704      Asn Arg Asp Trp Glu      Leu      Arg Gln Leu Phe
E--> 2705      Lys      Glu Glu          Met      Ala
E--> 2706
E--> 2707          35          40          45
      2709 Cys Phe Ser Lys Lys Thr Gly Val Ala Thr Glu Ala Gly Asp Thr Asn
E--> 2710      Val Ala Arg Arg Ala Ile Leu Val Ala Ala Ser      Glu Ile Glu
E--> 2711      Ile Phe      Ala Leu Glu Ile Ile Asp          Val Val
E--> 2712      Leu      Asn      Glu          Phe Gln
E--> 2713
E--> 2714          50          55          60
      2716 Val Glu Val Leu Lys Ala Lys Leu Lys His Val Ala Ser ^^^ Asp Glu
E--> 2717 Ala Asp Thr Phe Arg Glu      Val Thr Arg Asn Thr Asn Asp Pro
E--> 2718 Leu      His Ile      Thr      Phe Arg Lys      Ser Asp Asn
E--> 2719          Glu          Glu His
E--> 2720          65          70          75
      2722 Glu Val Asp Lys Ile Val Gln Lys Cys Val Val Lys Lys Ala Thr Pro
E--> 2723 Lys Ser Glu Asp Leu Ile Glu      Ala      Thr Glu Asp      Val
E--> 2724      Thr          Asn      Thr      Arg
E--> 2725          Ala
E--> 2726          80          85          90
      2728 Glu Glu Thr Ala Tyr Asp Thr Phe Lys Cys Ile Tyr Asp Ser Lys Pro
E--> 2729 Gln Asp Ser Val Phe Glu Val Thr      Val Val Leu Lys Asn Arg Ser
E--> 2730      His      Ser Ala Asn Phe          Met      Asp
E--> 2731
E--> 2732 95          100          105          110
      2734 Asp Phe Ser Pro Ile Asp ^^^ ^^^ *
E--> 2735 Asn      Phe Gly Asp Leu Phe Val * delete
E--> 2736 Lys      Val
E--> 2737          115
E--> 2740 221
E--> 2742 RB125 SEQ delete at end of file

```

*Invalid
grouping
of amino
acids*

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/876,796

DATE: 06/27/2001

TIME: 15:33:39

Input Set : A:\SEQLIST.txt

Output Set: N:\CRF3\06272001\I876796.raw

L:3 M:220 C: Keyword misspelled or invalid format, [(1) GENERAL INFORMATION:]
 L:8 M:220 C: Keyword misspelled or invalid format, [(iv) CORRESPONDENCE ADDRESS:]
 L:13 M:220 C: Keyword misspelled or invalid format, [(F) ZIP:]
 L:14 M:220 C: Keyword misspelled or invalid format, [(v) COMPUTER READABLE FORM:]
 L:20 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
 L:21 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]
 L:23 M:220 C: Keyword misspelled or invalid format, [(vii) PRIOR APPLICATION DATA:]
 L:26 M:220 C: Keyword misspelled or invalid format, [(viii) ATTORNEY/AGENT INFORMATION:]
 L:30 M:220 C: Keyword misspelled or invalid format, [(ix) TELECOMMUNICATION INFORMATION:]
 L:35 M:220 C: Keyword misspelled or invalid format, [(2) INFORMATION FOR SEQ ID NO:]
 L:47 M:220 C: Keyword misspelled or invalid format, [(C) INDIVIDUAL ISOLATE:]
 L:48 M:220 C: Keyword misspelled or invalid format, [(G) CELL TYPE:]
 L:52 M:220 C: Keyword misspelled or invalid format, [(ix) FEATURE:]
 L:60 M:220 C: Keyword misspelled or invalid format, [(2) INFORMATION FOR SEQ ID NO:]
 L:71 M:220 C: Keyword misspelled or invalid format, [(C) INDIVIDUAL ISOLATE:]
 L:72 M:220 C: Keyword misspelled or invalid format, [(G) CELL TYPE:]
 L:80 M:254 E: No. of Bases conflict, Input:36 Counted:46 SEQ:2
 M:254 Repeated in SeqNo=2
 L:124 M:204 E: No. of Bases differ, LENGTH:Input:566 Counted:576 SEQ:2
 L:127 M:220 C: Keyword misspelled or invalid format, [(2) INFORMATION FOR SEQ ID NO:]
 L:138 M:220 C: Keyword misspelled or invalid format, [(C) INDIVIDUAL ISOLATE:]
 L:139 M:220 C: Keyword misspelled or invalid format, [(G) CELL TYPE:]
 L:151 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:3
 L:171 M:342 E: Invalid Stop Code On Error, STOP CODON:*
 L:172 M:203 E: No. of Seq. differs, LENGTH:Input:134 Found:135 SEQ:3
 L:175 M:220 C: Keyword misspelled or invalid format, [(2) INFORMATION FOR SEQ ID NO:]
 L:186 M:220 C: Keyword misspelled or invalid format, [(C) INDIVIDUAL ISOLATE:]
 L:187 M:220 C: Keyword misspelled or invalid format, [(G) CELL TYPE:]
 L:216 M:342 E: Invalid Stop Code On Error, STOP CODON:*
 L:217 M:203 E: No. of Seq. differs, LENGTH:Input:116 Found:117 SEQ:4
 L:220 M:220 C: Keyword misspelled or invalid format, [(2) INFORMATION FOR SEQ ID NO:]
 L:231 M:220 C: Keyword misspelled or invalid format, [(C) INDIVIDUAL ISOLATE:]
 L:232 M:220 C: Keyword misspelled or invalid format, [(G) CELL TYPE:]
 L:282 M:220 C: Keyword misspelled or invalid format, [(2) INFORMATION FOR SEQ ID NO:]
 L:293 M:220 C: Keyword misspelled or invalid format, [(C) INDIVIDUAL ISOLATE:]
 L:294 M:220 C: Keyword misspelled or invalid format, [(G) CELL TYPE:]
 L:344 M:220 C: Keyword misspelled or invalid format, [(2) INFORMATION FOR SEQ ID NO:]
 L:355 M:220 C: Keyword misspelled or invalid format, [(C) INDIVIDUAL ISOLATE:]
 L:356 M:220 C: Keyword misspelled or invalid format, [(G) CELL TYPE:]
 L:388 M:342 E: Invalid Stop Code On Error, STOP CODON:*
 L:389 M:203 E: No. of Seq. differs, LENGTH:Input:133 Found:134 SEQ:7
 L:392 M:220 C: Keyword misspelled or invalid format, [(2) INFORMATION FOR SEQ ID NO:]
 L:403 M:220 C: Keyword misspelled or invalid format, [(C) INDIVIDUAL ISOLATE:]
 L:404 M:220 C: Keyword misspelled or invalid format, [(G) CELL TYPE:]
 L:433 M:342 E: Invalid Stop Code On Error, STOP CODON:*
 L:434 M:203 E: No. of Seq. differs, LENGTH:Input:115 Found:116 SEQ:8
 L:437 M:220 C: Keyword misspelled or invalid format, [(2) INFORMATION FOR SEQ ID NO:]
 L:448 M:220 C: Keyword misspelled or invalid format, [(C) INDIVIDUAL ISOLATE:]

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/876,796

DATE: 06/27/2001

TIME: 15:33:39

Input Set : A:\SEQLIST.txt

Output Set: N:\CRF3\06272001\I876796.raw

L:449 M:220 C: Keyword misspelled or invalid format, [(G) CELL TYPE:]
L:500 M:220 C: Keyword misspelled or invalid format, [(2) INFORMATION FOR SEQ ID NO:]
L:511 M:220 C: Keyword misspelled or invalid format, [(C) INDIVIDUAL ISOLATE:]
L:512 M:220 C: Keyword misspelled or invalid format, [(G) CELL TYPE:]
L:544 M:342 E: Invalid Stop Code On Error, STOP CODON:*
L:545 M:203 E: No. of Seq. differs, LENGTH:Input:133 Found:134 SEQ:10
L:548 M:220 C: Keyword misspelled or invalid format, [(2) INFORMATION FOR SEQ ID NO:]
L:559 M:220 C: Keyword misspelled or invalid format, [(C) INDIVIDUAL ISOLATE:]
L:560 M:220 C: Keyword misspelled or invalid format, [(G) CELL TYPE:]
L:589 M:342 E: Invalid Stop Code On Error, STOP CODON:*
L:590 M:203 E: No. of Seq. differs, LENGTH:Input:115 Found:116 SEQ:11
L:593 M:220 C: Keyword misspelled or invalid format, [(2) INFORMATION FOR SEQ ID NO:]
L:604 M:220 C: Keyword misspelled or invalid format, [(C) INDIVIDUAL ISOLATE:]
L:605 M:220 C: Keyword misspelled or invalid format, [(G) CELL TYPE:]
L:656 M:220 C: Keyword misspelled or invalid format, [(2) INFORMATION FOR SEQ ID NO:]
L:667 M:220 C: Keyword misspelled or invalid format, [(C) INDIVIDUAL ISOLATE:]
L:668 M:220 C: Keyword misspelled or invalid format, [(G) CELL TYPE:]
L:691 M:333 E: Wrong sequence grouping, Amino acids not in groups!
L:692 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:13
M:332 Repeated in SeqNo=13
L:700 M:342 E: Invalid Stop Code On Error, STOP CODON:*
L:704 M:220 C: Keyword misspelled or invalid format, [(2) INFORMATION FOR SEQ ID NO:]
L:736 M:333 E: Wrong sequence grouping, Amino acids not in groups!
L:737 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:14
M:332 Repeated in SeqNo=14
L:745 M:342 E: Invalid Stop Code On Error, STOP CODON:*
L:773 M:254 E: No. of Bases conflict, Input:90 Counted:91 SEQ:15
L:938 M:342 E: Invalid Stop Code On Error, STOP CODON:*
L:939 M:203 E: No. of Seq. differs, LENGTH:Input:169 Found:174 SEQ:17
L:1056 M:342 E: Invalid Stop Code On Error, STOP CODON:*
L:1057 M:203 E: No. of Seq. differs, LENGTH:Input:149 Found:150 SEQ:19
L:1186 M:342 E: Invalid Stop Code On Error, STOP CODON:*
L:1187 M:203 E: No. of Seq. differs, LENGTH:Input:169 Found:174 SEQ:21
L:1304 M:342 E: Invalid Stop Code On Error, STOP CODON:*
L:1305 M:203 E: No. of Seq. differs, LENGTH:Input:149 Found:150 SEQ:23
L:1376 M:254 E: No. of Bases conflict, Input:595 Counted:591 SEQ:24
M:254 Repeated in SeqNo=24
L:1388 M:204 E: No. of Bases differ, LENGTH:Input:777 Counted:776 SEQ:24
L:1421 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:25
M:342 Repeated in SeqNo=25
L:1442 M:203 E: No. of Seq. differs, LENGTH:Input:170 Found:175 SEQ:25
L:1559 M:342 E: Invalid Stop Code On Error, STOP CODON:*
L:1560 M:203 E: No. of Seq. differs, LENGTH:Input:149 Found:150 SEQ:27
L:1689 M:342 E: Invalid Stop Code On Error, STOP CODON:*
L:1690 M:203 E: No. of Seq. differs, LENGTH:Input:173 Found:174 SEQ:29
L:1750 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:30
L:1754 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:30
L:1758 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:30
L:1808 M:342 E: Invalid Stop Code On Error, STOP CODON:*

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/876,796

DATE: 06/27/2001

TIME: 15:33:39

Input Set : A:\SEQLIST.txt

Output Set: N:\CRF3\06272001\I876796.raw

L:1809 M:203 E: No. of Seq. differs, LENGTH:Input:149 Found:150 SEQ:31
L:1932 M:333 E: Wrong sequence grouping, Amino acids not in groups!
L:1933 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:33
M:332 Repeated in SeqNo=33
L:1938 M:342 E: Invalid Stop Code On Error, STOP CODON:*
L:2047 M:333 E: Wrong sequence grouping, Amino acids not in groups!
L:2048 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:35
M:332 Repeated in SeqNo=35
L:2056 M:342 E: Invalid Stop Code On Error, STOP CODON:*
L:2101 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:36
L:2105 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:36
L:2110 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:36
L:2114 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:36
L:2118 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:36
L:2122 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:36
L:2126 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:36
L:2130 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:36
L:2189 M:342 E: Invalid Stop Code On Error, STOP CODON:*
L:2190 M:203 E: No. of Seq. differs, LENGTH:Input:173 Found:174 SEQ:37
L:2307 M:342 E: Invalid Stop Code On Error, STOP CODON:*
L:2308 M:203 E: No. of Seq. differs, LENGTH:Input:149 Found:150 SEQ:39
L:2528 M:204 E: No. of Bases differ, LENGTH:Input:481 Counted:484 SEQ:45
L:2594 M:204 E: No. of Bases differ, LENGTH:Input:481 Counted:484 SEQ:46
L:2660 M:204 E: No. of Bases differ, LENGTH:Input:481 Counted:484 SEQ:47
L:2687 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:0
M:332 Repeated in SeqNo=48
L:2716 M:330 E: (2) Invalid Amino Acid Designator, 1
L:2734 M:342 E: Invalid Stop Code On Error, STOP CODON:*
L:2734 M:330 E: (2) Invalid Amino Acid Designator, 2
M:342 Repeated in SeqNo=48
L:2742 M:333 E: Wrong sequence grouping, Amino acids not in groups!
L:2742 M:330 E: (2) Invalid Amino Acid Designator, 2
L:2742 M:203 E: No. of Seq. differs, LENGTH:Input:133 Found:320 SEQ:48